–Supplementary File–
Dec-Adapter: Exploring Efficient Decoder-side Adapter for Bridging Screen Content and Natural Image Compression

Sheng Shen, Huanjing Yue*, Jingyu Yang*
Tianjin University
{codyshens, huanjing.yue, yjy}@tju.edu.cn

1. Experiment Settings for Compared Methods

In this section, we describe the experimental setup for the five existing adaptive compression methods. Yang *et al.* [7] perform only latent refinement without tuning decoders, which corresponds to the first stage of our approach. Lam *et al.* [4] update the bias parameters in the convolutional layers of the decoder after latent refinement. These parameters are optimized solely for distortion and are converted to 64-bit floating-point numbers before being compressed using 7z format. Therefore, we update it in the second stage with only distortion.

Rozendaal *et al.* [6] update all parameters in both the decoder and entropy model and optimize them for rate-distortion trade-off. Therefore, we optimize them directly for rate-distortion.

Zou *et al.* [8] use overfittable multiplicative parameters (OMPs) instead of adapters after latent refinement and optimize them solely for distortion. The updated parameters are linearly transformed to a range between 0 and 255 before being quantized to integers to obtain eight-bit integer values and two real values (scale and bias) for linear transformation with a precision of 32 bits each. Therefore, we optimize it in the second stage solely for distortion.

Finally, Tsubota *et al.* [5] use matrices as learnable adapters after latent refinement and optimize them for both rate and distortion. They further adopt matrix decomposition to reduce the number of parameters. Therefore, we update it in the second stage with both rate and distortion.

2. Adapted to Other Domains

Our adapter can also be applied to images in other domains. Specifically, we evaluate on three additional domains: Comic, Line drawing, and Vector art, which are collected by [5] by utilizing WACNN [9] as the baseline. The results are presented in Table 1. For Kodak, which is an in-domain dataset, the bit-reduction brought by the adapter is smaller than the bit-increment brought by its parameters. Yang *et al.*'s method optimizes the latent code without extra bitstreams, yielding superior results compared to ours. Nevertheless, in most cases, our proposed method outperforms the other adaptation methods, particularly when the image contains a large portion of synthetic content that are far from natural content. It should be noted that on the Kodak dataset, which is an in-domain test, our results are slightly lower than the baseline but still superior to VVC [2]. This differs from the case where Cheng2020 [1] is used as a baseline. This is because WACNN has higher compression performance in the natural domain; when an adapter is inserted, performance improvement is limited or even non-existent (the additional overhead bits outweigh the improvement in PSNR value). Figure 1 presents the results of our rate-distortion (RD) performance comparison. Our Dec-Adapter outperforms other learned compression methods when the image contains rich information and is different from the previous domain.

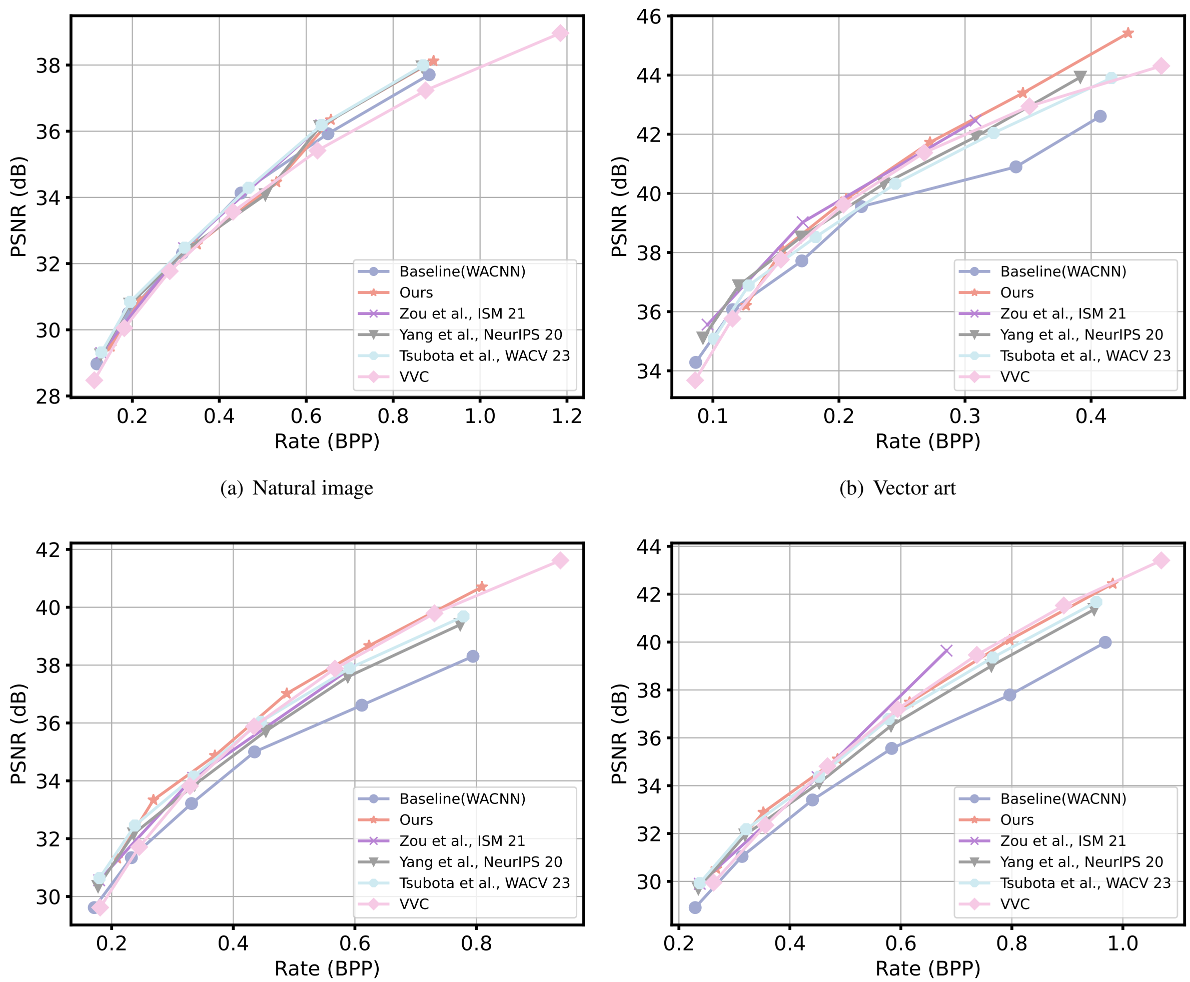
3. Qualitative Results

We present additional visual qualitative results comparing our method with other compression methods in Figure 2 and 3. All the results are generated with WACNN [9] as the baseline. The figures show examples of reconstructed images using our method and those of Yang *et al.* [7], Rozendaal *et al.* [6], Zou *et al.* [8], Lam *et al.* [4], and Tsubota *et al.* [5]. Our reconstructed images retain more detail at similar bit rates.

*Corresponding authors. This work was supported in part by the National Natural Science Foundation of China under Grant 62231018 and Grant 62072331.

Table 1. Comparison with existing adaptive compression methods in terms of BD rate (%) over WACNN [9]. A smaller value means more effective.

Method	Description	Kodak (In Domain)	Vector (Out Domain)	Line (Out Domain)	Comic (Out Domain)	Average
VVC [2]	H.266	6.72	-9.41	-10.41	-12.69	-6.45
WACNN [9]	Baseline	0.00	0.00	0.00	0.00	0.00
Yang *et al.* [7]	Only refine latent	0.31	-13.49	-11.57	-10.81	-8.89
Lam *et al.* [4]	Bias-tuning	197	1101	173	138	402
Rozendaal *et al.* [6]	Full para. fine-tuning	335	685	323	259	400
Zou *et al.* [8]	OMP-tuning	-1.70	-19.37	-13.83	-14.32	-12.30
Tsubota *et al.* [5]	MD-tuning	-3.74	-8.62	-16.09	-13.69	-10.53
Ours	Conv-tuning	2.84	-16.05	-18.90	-14.86	-11.74



(a) Natural image

(b) Vector art

(c) Comic

(d) Line drawing

Figure 1. Comparison with attention based compression method WACNN [9], which is the baseline method that does not perform adaptive optimization.

References

[1] Z. Cheng, H. Sun, M. Takeuchi, and J. Katto. Learned image compression with discretized gaussian mixture likelihoods and attention modules. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7939–7948, 2020. 1

[2] H. de Jesús Ochoa Domínguez and K. R. Rao. Versatile video coding. 2019. 1, 2

[3] E. Kodak. Kodak lossless true color image suite (photocd pcd0992). *URL http://r0k. us/graphics/kodak*, 6, 1993. 4

[4] Y.-H. Lam, A. Zare, F. Cricri, J. Lainema, and M. M. Hannuksela. Efficient adaptation of neural network filter for video

compression. In *Proceedings of the 28th ACM International Conference on Multimedia*, pages 358–366, 2020. 1, 2

[5] K. Tsubota, H. Akutsu, and K. Aizawa. Universal deep image compression via content-adaptive optimization with adapters. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 2529–2538, 2023. 1, 2, 5

[6] T. van Rozendaal, I. A. Huijben, and T. S. Cohen. Overfitting for fun and profit: Instance-adaptive data compression. *arXiv preprint arXiv:2101.08687*, 2021. 1, 2

[7] Y. Yang, R. Bamler, and S. Mandt. Improving inference for neural image compression. *Advances in Neural Information Processing Systems*, 33:573–584, 2020. 1, 2

[8] N. Zou, H. Zhang, F. Cricri, R. G. Youvalari, H. R. Tavakoli, J. Lainema, E. Aksu, M. Hannuksela, and E. Rahtu. Adaptation and attention for neural video coding. In *2021 IEEE International Symposium on Multimedia (ISM)*, pages 240–244. IEEE, 2021. 1, 2

[9] R. Zou, C. Song, and Z. Zhang. The devil is in the details: Window-based attention for image compression. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17492–17501, 2022. 1, 2

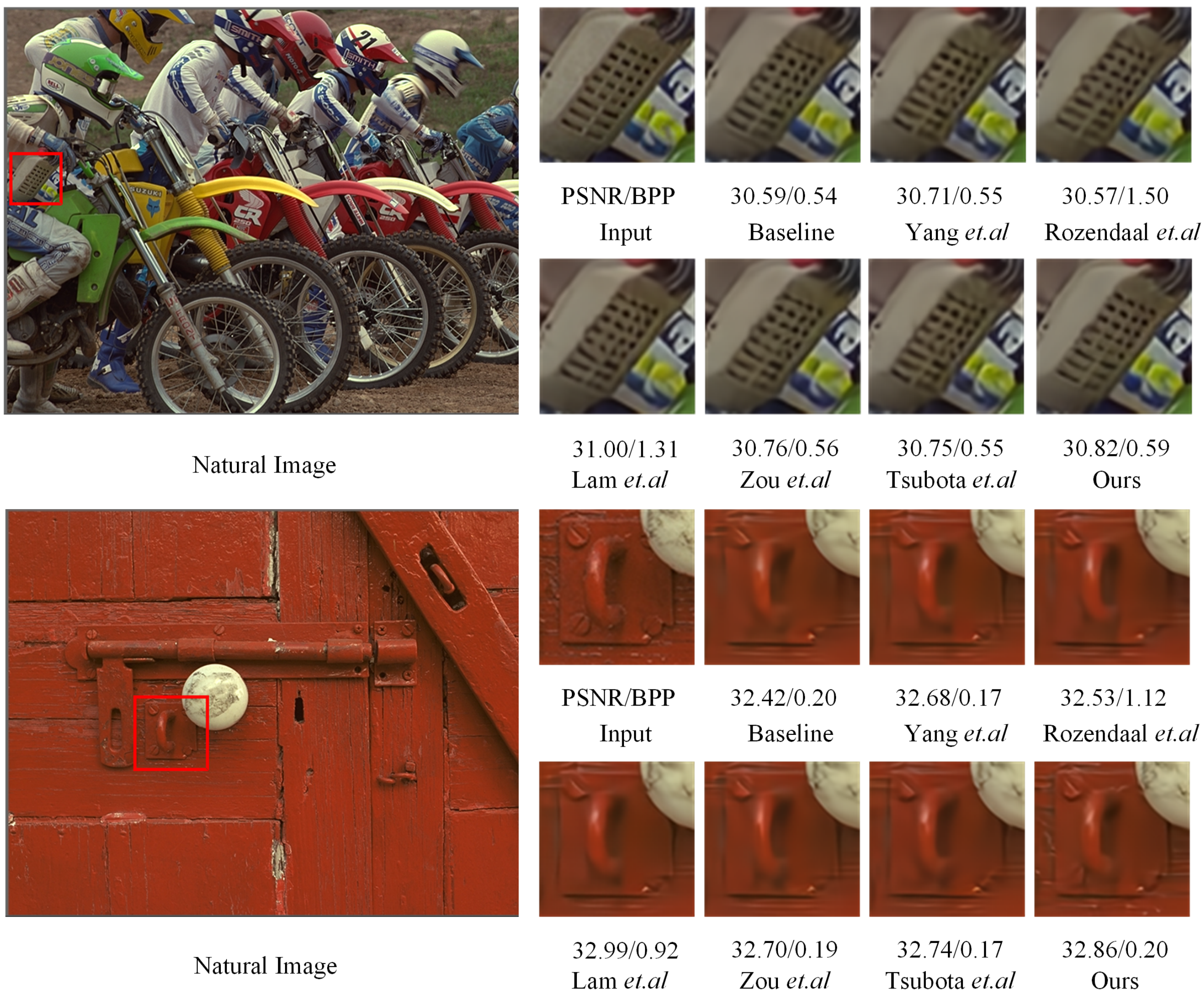


Figure 2. Qualitative results for different adaptive compression methods on natural dataset Kodak [3].

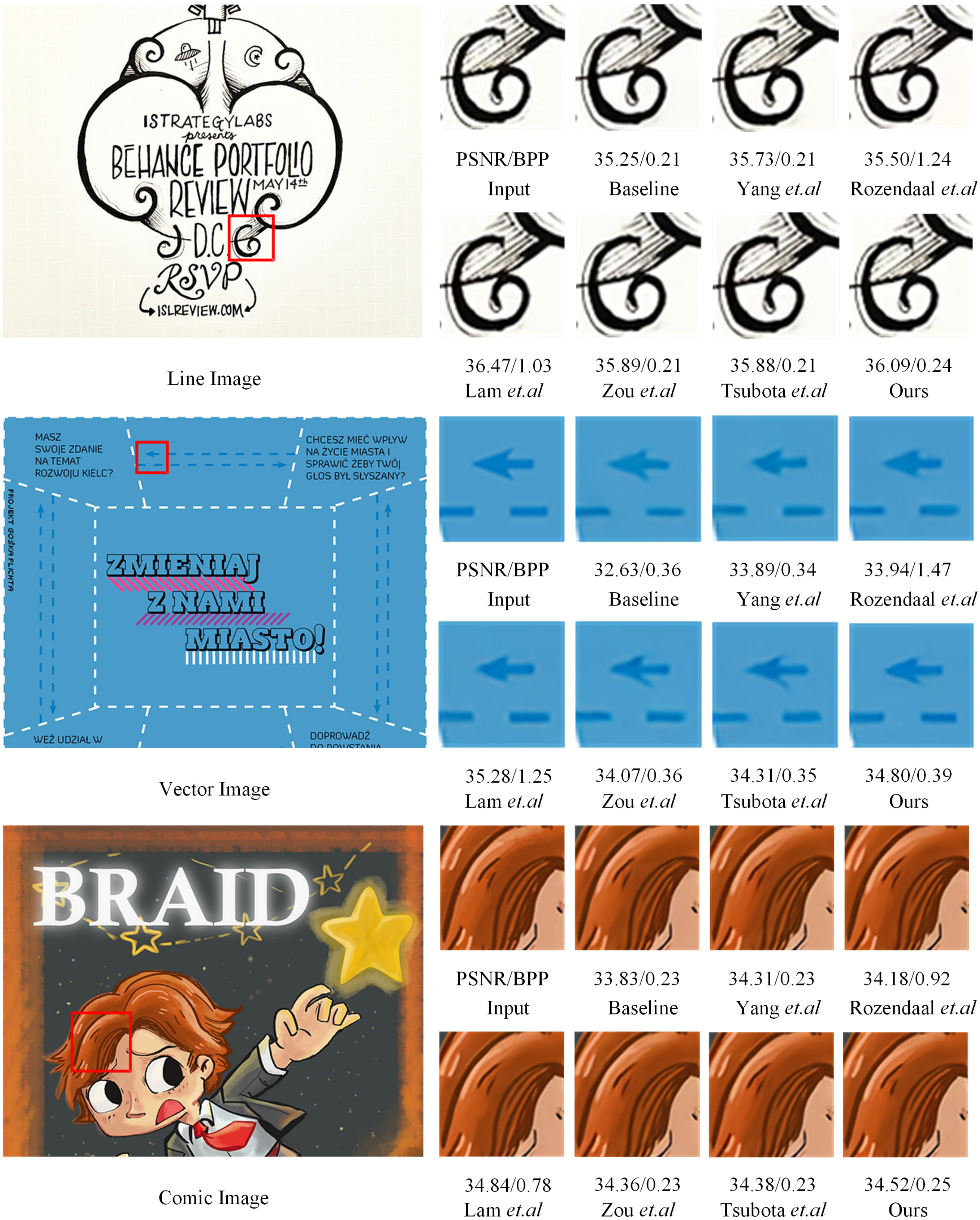


Figure 3. Qualitative results for different adaptive compression methods on other three domains [5].